

- 27 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: (OTHER THAN US): THE UNIVERSITY OF MELBOURNE  
(US ONLY): SINGH Mohan, BHALLA Prem, HUI-LING Xu and  
SWOBODA Ines

(ii) TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
THEREFOR

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE  
(B) STREET: 1 LITTLE COLLINS STREET  
(C) CITY: MELBOURNE  
(D) STATE: VICTORIA  
(E) COUNTRY: AUSTRALIA  
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT INTERNATIONAL  
(B) FILING DATE: 24-JUL-1998  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO8233  
(B) FILING DATE: 25-JUL-1997  
(C) CLASSIFICATION:

(viii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PP1184  
(B) FILING DATE: 31-DEC-1997  
(C) CLASSIFICATION:

(ix) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES, DR E JOHN L  
(C) REFERENCE/DOCKET NUMBER: EJH/AF

- 28 -

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

(C) TELEX: AA 31787

- 29 -

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTACTCTTAA GCATACAACA TGAG

14

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAGGCATACT TGAATGCTAC AAGA

14

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 625 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 82..468

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCATCCCAT CAACAGAAGG TTTAAGTGGA AATCCATTTT ATTAGAAAAG ATCGGACAAA	60
GGGTACTCTT AAGCATACAA C ATG AGG GCG GTG GCG GTT TTC TTT GCT TGC	111
Met Arg Ala Val Ala Val Phe Phe Ala Cys	
1 5 10	
GTT CTC TTC TGT ATG GTT CAC AAA GCC GCA CTT GCG GAT GAT AAA ACG	159
Val Leu Phe Cys Met Val His Lys Ala Ala Leu Ala Asp Asp Lys Thr	
15 20 25	
TGC AAC CCT ACA GAT TTT ATG GTT ACC CAA ACC ATA ACT GGA TTG ACA	207

- 30 -

Cys	Asn	Pro	Thr	Asp	Phe	Met	Val	Thr	Gln	Thr	Ile	Thr	Gly	Leu	Thr		
			30					35					40				
ATC	GGC	GGT	AAA	CAA	GAG	TTC	GAG	GTC	AAT	TTA	ATA	AAC	AAT	TTG	TAT		255
Ile	Gly	Gly	Lys	Gln	Glu	Phe	Glu	Val	Asn	Leu	Ile	Asn	Asn	Leu	Tyr		
		45					50					55					
TGT	GCA	CAA	TCT	AAT	GTC	AAA	GTT	TCA	TGT	GAC	GGG	CTT	CAT	ACC	ACC		303
Cys	Ala	Gln	Ser	Asn	Val	Lys	Val	Ser	Cys	Asp	Gly	Leu	His	Thr	Thr		
	60					65					70						
GAA	CCA	ATA	GAT	CCT	CAC	ATT	ATC	AGA	CCA	CTT	AGT	GAC	GGA	ACG	AAC		351
Glu	Pro	Ile	Asp	Pro	His	Ile	Ile	Arg	Pro	Leu	Ser	Asp	Gly	Thr	Asn		
	75				80					85					90		
AAC	TGC	CTT	GTC	AAC	AAT	GGA	GCG	CCT	ATT	TCT	CAT	GCT	ACT	CTT	GTA		399
Asn	Cys	Leu	Val	Asn	Asn	Gly	Ala	Pro	Ile	Ser	His	Ala	Thr	Leu	Val		
				95				100					105				
GCA	TTC	AAG	TAT	GCC	TGG	GAT	GTT	CCT	CCA	TCT	TTC	AGC	ATC	ATC	AGC		447
Ala	Phe	Lys	Tyr	Ala	Trp	Asp	Val	Pro	Pro	Ser	Phe	Ser	Ile	Ile	Ser		
			110				115						120				
TCT	GAT	ATA	AAT	TGC	TCC	TAA	GGAGAAA	ATTCTAGTTG	GCAGAGAATA								495
Ser	Asp	Ile	Asn	Cys	Ser	OCH											
		125															
ATCATATAGT	CTTTTTTACT	GAGCTATTTA	ATTTTTTCAA	TTTTCACCAA	TAAGATTATT												555
TTAATGGAAT	GTTAATGTAT	TAGAATTGAA	AAATAAAAAA	AAAAAAAAAA	AAAAAAAAAA												615
AAAAAAAAAA																	625

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 128 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Arg	Ala	Val	Ala	Val	Phe	Phe	Ala	Cys	Val	Leu	Phe	Cys	Met	Val		
1				5					10					15			
His	Lys	Ala	Ala	Leu	Ala	Asp	Asp	Lys	Thr	Cys	Asn	Pro	Thr	Asp	Phe		
		20						25					30				
Met	Val	Thr	Gln	Thr	Ile	Thr	Gly	Leu	Thr	Ile	Gly	Gly	Lys	Gln	Glu		
		35					40					45					
Phe	Glu	Val	Asn	Leu	Ile	Asn	Asn	Leu	Tyr	Cys	Ala	Gln	Ser	Asn	Val		
	50					55					60						
Lys	Val	Ser	Cys	Asp	Gly	Leu	His	Thr	Thr	Glu	Pro	Ile	Asp	Pro	His		
	65				70				75					80			
Ile	Ile	Arg	Pro	Leu	Ser	Asp	Gly	Thr	Asn	Asn	Cys	Leu	Val	Asn	Asn		
			85					90						95			
Gly	Ala	Pro	Ile	Ser	His	Ala	Thr	Leu	Val	Ala	Phe	Lys	Tyr	Ala	Trp		
		100					105						110				
Asp	Val	Pro	Pro	Ser	Phe	Ser	Ile	Ile	Ser	Ser	Asp	Ile	Asn	Cys	Ser	OCH	

- 31 -

115

120

125

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 587 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 49..378

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5.

GAAAGTTGAA ACATCTCCAT CAAACTCTAG AGTCAGATTT CCCACAAG ATG ATT TCA	57
Met Ile Ser	
1	
TCG GCA AAT AAC AAA GGC GCC GGC ACA AGC CGC CGC AAG CTC CGT TCT	105
Ser Ala Asn Asn Lys Gly Ala Gly Thr Ser Arg Arg Lys Leu Arg Ser	
5 10 15	
GAG AAG GCT GCA CTC CAG TTC TCC GTC AGT CGC GTC GAA TAC TCC CTC	153
Glu Lys Ala Ala Leu Gln Phe Ser Val Ser Arg Val Glu Tyr Ser Leu	
20 25 30 35	
AAG AAG GGG CGC TAT TGC AGG CGC TTA GGC GCT ACG GCC CCC GTC TAC	201
Lys Lys Gly Arg Tyr Cys Arg Arg Leu Gly Ala Thr Ala Pro Val Tyr	
40 45 50	
CTA GCC GCC GTC CTT GAA AAC CTC GTG GCC GAA GTG TTG GAC ATG GCG	249
Leu Ala Ala Val Leu Glu Asn Leu Val Ala Glu Val Leu Asp Met Ala	
55 60 65	
GCG AAC GTG ACA GAA GAA ACA TCC CCC ATT GTT ATC AAA CCG AGG CAT	297
Ala Asn Val Thr Glu Glu Thr Ser Pro Ile Val Ile Lys Pro Arg His	
70 75 80	
ATT ATG CTT GCC CCC AGG AAT GAT GTA GAA GTT GAA CAA GCT GTT TCA	345
Ile Met Leu Ala Pro Arg Asn Asp Val Glu Val Glu Gln Ala Val Ser	
85 90 95	
CGG TGT CAC CAT CTC GGC ATC AGG TGT CGT CCC TAAACACGC AAAGAGCTGG	398
Arg Cys His His Leu Gly Ile Arg Cys Arg Pro	
100 105 110	
ACCGTCGCAA ACGCCGTTCC ACCTTTCAGC CGGATTAGTT CTTGATATTT CATTCTATCA	458
ATCTTGTTTA TGTGACTGTG ATTTTTCGTT TTGTGTTGAA CTAAGCCCCC TAATCTGGAT	518
TTCTCGTTTT ATGTTGAACT AAGTCTGTGC ACTCTTGAAG TAAAAAAAAA AAAAAAAAAA	578
AAAAAAAAA	587

## (2) INFORMATION FOR SEQ ID NO:6:

- 32 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Ile Ser Ser Ala Asn Asn Lys Gly Ala Gly Thr Ser Arg Arg Lys
 1           5           10           15
Leu Arg Ser Glu Lys Ala Ala Leu Gln Phe Ser Val Ser Arg Val Glu
          20           25           30
Tyr Ser Leu Lys Lys Gly Arg Tyr Cys Arg Arg Leu Gly Ala Thr Ala
          35           40           45
Pro Val Tyr Leu Ala Ala Val Leu Glu Asn Leu Val Ala Glu Val Leu
          50           55           60
Asp Met Ala Ala Asn Val Thr Glu Glu Thr Ser Pro Ile Val Ile Lys
 65           70           75           80
Pro Arg His Ile Met Leu Ala Pro Arg Asn Asp Val Glu Val Glu Gln
          85           90           95
Ala Val Ser Arg Cys His His Leu Gly Ile Arg Cys Arg Pro
          100          105          110

```

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 485 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 16..348

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

GATCCCAAAT CATCA ATG ACG ATC CCC GAA AAG AAA TCC GTC GCT CCG ATG      51
      Met Thr Ile Pro Glu Lys Lys Ser Val Ala Pro Met
          1           5           10
GCC CGT ATG AAG CAT ACA GCC CGC ATG TCT ACC GGC GGT AAG GCT CCA      99
Ala Arg Met Lys His Thr Ala Arg Met Ser Thr Gly Gly Lys Ala Pro
          15           20           25
CGC AAG CAG CTC GCC TCT AAG GCT CTT CGC AAG GCG CCA CCA CCA CCG     147
Arg Lys Gln Leu Ala Ser Lys Ala Leu Arg Lys Ala Pro Pro Pro Pro
          30           35           40
ACC AAA GGA GTG AAG CAG CCC ACC ACT ACC ACC TCC GGA AAA TGG CGC     195
Thr Lys Gly Val Lys Gln Pro Thr Thr Thr Thr Ser Gly Lys Trp Arg
          45           50           55           60
TTC GCG AGA TTT CAC AGG AAA CTG CCA TTC CAA GGG CTG GTG AGG AAA     243
Phe Ala Arg Phe His Arg Lys Leu Pro Phe Gln Gly Leu Val Arg Lys
          65           70           75

```

- 33 -

ATC TGG CAG GAC TTG AAG ACA CAT CTG CGC TTC AAG AAC CAC TCG GTT	291
Ile Trp Gln Asp Leu Lys Thr His Leu Arg Phe Lys Asn His Ser Val	
80 85 90	
CCT CCA CTT GAG GAG GTA ACT GAG GTT TAT CCT TGC CAA ACT ATT GGA	339
Pro Pro Leu Glu Glu Val Thr Glu Val Tyr Pro Cys Gln Thr Ile Gly	
95 100 105	
GGA TGC TAT TAGGATATTG AATTGGATA ATGGTTTAAT TATCTGTTCT	388
Gly Cys Tyr	
110	
ACCTTTATGA TCAAATTTCT GTGGCTCAGC GTTGTGTAAT TTGGGCAATC GAATTCTTAG	448
CTATATTGCC TCAAAAAAAA AAAAAAAAAA AAAAAAA	485

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 111 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Thr Ile Pro Glu Lys Lys Ser Val Ala Pro Met Ala Arg Met Lys	
1 5 10 15	
His Thr Ala Arg Met Ser Thr Gly Gly Lys Ala Pro Arg Lys Gln Leu	
20 25 30	
Ala Ser Lys Ala Leu Arg Lys Ala Pro Pro Pro Pro Thr Lys Gly Val	
35 40 45	
Lys Gln Pro Thr Thr Thr Thr Ser Gly Lys Trp Arg Phe Ala Arg Phe	
50 55 60	
His Arg Lys Leu Pro Phe Gln Gly Leu Val Arg Lys Ile Trp Gln Asp	
65 70 75 80	
Leu Lys Thr His Leu Arg Phe Lys Asn His Ser Val Pro Pro Leu Glu	
85 90 95	
Glu Val Thr Glu Val Tyr Pro Cys Gln Thr Ile Gly Gly Cys Tyr	
100 105 110	

- 34 -

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 945 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGAGGGTGTT GGAATTAGGT TTGCCTAGGG TTTGCCTAGG TTTAGAGAAA TAGTCAAAAT	60
TGTCCTATTC TATAGGCATG ATTTAGTAGT GAGTTAATTA TCCTATAATT TCTCTTCTTG	120
TATGCTCAAA TAACTGGTTC TTTAATGAAT AGATAATTAA GTTTTGTAGC AATTTCTTCC	180
TCAAATTGAG TATCAACAAT TGTTAGATTG CTTTGGTGAT TATATTTGAT ATAATTGTTT	240
GTAAGAATGT GTAGTGAAAA GATTGTGATT ATTCATTTCC TTGTTGGACG AATTGTTAGA	300
GCCCCATCGC TAATGCCTTA TAGTACTCGA AATATGTTGG GAATAGAAGA TGAAAAATCC	360
CATTCTTTGT AGTAGGAGTA AAAATTTGTC TTTTCATTAT TCCATTGAAT GTTAACCACT	420
TGCCATTCAT CTGACGGGGA TGGCAGAGTT CCGACCATCT AGTGATCCGT GGGATATTGA	480
TTTTGGTGTG TCAATGAAAT TGTGAGAACG GGCTTCTGGG AGAGAAAAGC CCTCTTGCCT	540
CTGATATGAA CACTGAGGCT GATTATGTTA ACGGATGGAG ATTTATCAGT GGCTGAATTT	600
GGGTGCTGTA GAGACAGAAT TTGAAAGTTC TAACAATAAA CCCTAATTCT GAACTTGGGC	660
GGGGCTGGGA TTTTACTCTT AACGTGAAGA GAGGCAAGAT GAATTGACAG CTTGGAAGTC	720
GATCCAGTAT TTGCAGCAGT CGTGACGAAT TGGTTGGACA GTTACATCGG TCAGAGAATG	780
CGTTCTATAA ATTCCCCCAA TGCGGCAGTG AAAATCCCAT CCCATCAACA GAAGTTTTAA	840
GTGGAAACCC ATTCCAATAG AGAAGATCGA ACAAAGGGTA TTAAACATA CAAATGGGGG	900
CAGTGGTGTT TCTTTTGTCT TGCGTTCTCT TCTGTATGGT TCACA	945

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CAGGCATACT TGAATGCTAC AAGA

14

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs



- 35 -

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGTGAACCAT ACAGAAGAGA ACGC

24